Method for Stochastic Selection of Improved Cost Metric Backup Paths in Shared-Mesh Protection Networks

[001] **Background of the Invention**

[002] **Field of the Invention**

[003] This invention relates generally to the field of networking and in particular to methods of selecting and provisioning communications paths in shared-mesh protection communications networks.

[004] **Description of the Prior Art**

[005] Communications networks have been developed to provide communications services in the form of transmission of information carrying signals between ingress points and egress points of the network in a reliable and cost-efficient manner. Networks are comprised of a plurality of sets of communications equipment, or network elements, the bi-directional transmission links interconnecting them, and the software used to control signal propagation through the network. Each link is comprised of fibers and optical devices that enable the propagation between adjacent network elements of concurrent, uncorrelated optical channels. Signaling is in turn achieved by pulsing the channels at predetermined rates. Each signal traverses the network from source client equipment connected to its ingress network element to destination client equipment connected to its egress network element over a primary path comprised of a serial chain of network elements and links carrying channels. Since the amount of equipment and number of links determine the cost of the network, primary paths are typically kept as short as possible while still providing desired connectivity.

[006] Unfortunately, network elements and transmission links, including those comprising primary paths, exhibit a variety of failure modes including equipment malfunction, software failure, and human failure that can result in loss of service between ingress and egress network elements. The impact of service interruption can be severe due to the amount of

information carried on each signal in a typical modern network. As such, it is desirable to protect service by providing an alternate path for each signal, called the back-up path, that is used to carry the signal in the event of a failure on the signal's primary path.

[007] One method to provide service protection is to dedicate a backup path for each primary path in the network. In such dedicated protection methods, the primary path and backup path both carry a copy of the signal at all times and the egress node continually monitors the signal and decides which copy to use on an on-going basis. Dedicated protection provides the fastest network restoration time but at high cost since the amount of equipment is at least twice that of an unprotected network.

[008] An alternative method of network protection is shared-mesh protection, wherein the network elements and channels comprising backup paths can be used to protect multiple primary paths simultaneously. In such shared-mesh protection methods, a backup path does not carry live traffic until one of the primary paths it is protecting incurs a failure. Once a failure on a protected primary paths occurs, the network elements and channels of the backup path begin carrying the affected signal and the other primary paths protected by links in this backup path become unprotected until the network is reprovisioned.

[009] The concept of Shared Risk Group (SRG) is described by Labourdette, Bouillet, Ramamurthy, EllinasChaudhuri, and Bala in a paper Routing Strategies for Capacity-Efficient and Fast Restorable Mesh Optical Networks," which appeared in Photonic Network Communications, vol. 4, in July 2002, on pages 219 through 235, and also in a paper by Ellinas, Bouillet, Ramamurthy, Labourdette, Chaudhuri, and Bala, in a paper entitled "Routing and Restoration Architectures in Mesh Optical Networks," in Optical Netwrok Magazine, vol. 4, in January 2003, on pages 91 through 106. As taught by these authors, each point of independent failure in a network is denoted by an SRG with an associated list that includes all paths that share that risk of failure. By way of example, a section of conduit containing links of multiple paths, multiple signal lines wrapped in a single cable, or multiple signal channels multiplexed onto a single signal line would be types of SRGs.

[010] In order to avoid having a single failure event require two or more primary paths to obtain signal restoration from the same shared protection channel on a backup path link, all primary paths protected by the same shared protection channel must be SRG disjoint. Two or more primary paths are described as SRG disjoint if the sets of SRGs associated with each path have no common elements. Therefore, all prospective channels on proposed backup path links must not have any SRGs in common with the primary path seeking protection, and the primary path seeking protection must not have any SRGs in common with any primary paths already being protected by the prospective channels on the proposed backup path links.

[011] Network control, the manner in which signals propagate through the network and the computation and provisioning of primary and backup paths, is provided by suites of network management software collectively known as the network controller. Control functionality can be affected either by means of a central control element that communicates with each network element directly or by distributed local control functions located at network element that communicate with each other on an as-needed basis.

[012] In the case of a central control element, the control element maintains a database of all state variables describing (a) each network element and all links interconnecting them such as a list of links between adjacent nodes, the cost of using those links, the list of SRGs each link traverses, the number of available channels and their data rates between adjacent nodes, the number of shared channels reserved for protection, and their data rates between adjacent nodes, and (b) the provisioned primary and backup paths and corresponding channels, and the union set of all the SRG's traversed by all links of all primaries protected by each shared protection channel. The network controller communicates with each individual network element in order to provision new services, respond to changes in the network such as failures, and to maintain and update its database.

[013] As the size of the network grows, the size and complexity of the database required to manage the network with a central controller becomes unmanageable. An alternative control method for large networks utilizing intelligent network elements and distributed control is

described by Ramamurthy, Sengupta, and Chaudhuri in paper entitled "Comparison of Centralized and Distributed Provisioning of Paths in Optical Networks," which appeared in The Technical Digest of the Optical Fiber Communication Conference, 2001 on pages MH4-1-3.

[014] In the alternative, distributed control approach, a signal's ingress network element controls the computation of primary and backup paths to its intended egress network element. In contrast to a central control method wherein complete SRG information about every link in the entire network is available to the controller, in the distributed control method the information held at each network element about the SRGs in the network is a subset of the complete information. Specifically, SRG information is limited to the number of times each SRG is traversed by a primary path being protected by each shared protection channel in the network. Specifically, each network element holds a local database of state variables that include the links to adjacent network elements, the cost of using each link prorated for each channel carried on it, a list of SRGs each link traverses, the number and data rates of available channels between each pair of network elements, the number of shared protection channels between adjacent network elements, and the number of times each SRG in the network is traversed by a primary path protected by any shared protection channel in that link, where the set of SRGs traversed by each primary is the union set of SRGs traversed by all the links along that path.

[015] In the distributed control method, the process used to compute primary and backup paths is an iterative one wherein the ingress network element designates a set of primary paths and subsequently requests each network element on those paths to propose a link to adjacent network elements on the paths. The proposed links are compared to the SRG of the primary and its SRGs. If commonality of SRGs is determined, the choice is invalid, and a request for another proposed link is made. This process continues until a valid link choice is identified or the list of potential candidates is exhausted.

[016] Unfortunately, this iterative process is generally unsatisfactory for shared-mesh protection networks since the resulting networks typically require more network elements and links than necessary. Since the selection of backup paths is effectively a random process, the likelihood of an optimized selection of primary and backup paths is extremely low. A consequence of the non-optimal configuration is that additional transmission equipment is required, thus higher capital costs and greater latency during switching.

[017] A need therefore exists of a method to compute sets of primary and backup channels in distributed-control, shared-mesh protection networks that is independent of network size.

[018] **Summary of the Invention**

[019] We have developed a method of path selection in shared-mesh restoration networks that results in efficient use of network resources, using only the network state information available locally at each network element. The method uses probability theory to develop an estimate of protection channel sharing opportunities and encourages sharing of protection channels if possible.

[020] Viewed from a first aspect, our invention is a method of determining a set of primary and backup paths in a communications network that makes efficient use of available network resources without requiring a centralized controller or large complicated databases of state variables. It is a stochastic method that uses probability theory to estimate the ability to share proposed backup channels for each given proposed primary path without incurring contention problems. Our method is described by Bouillet, Labourdette, Ellinas, Ramamurthy, and Chaudhuri in a paper entitled "Stochastic Approaches to Route Shared Mesh Restored Lightpaths in Optical Mesh Networks," in the Proceedings of the Conference on Computer Communications, in June 2002, on pages 801 through 807. Advantageously and according to our invention, backup links are efficiently shared thereby reducing the total amount of network resources, i.e. the cost, needed to attain a given capacity. Our inventive method employs signaling between network elements to provision the primary and backup paths

once the lowest cost candidates have been determined. According to our invention, a set of shortest paths from an ingress network element to an egress network element is chosen as a candidate group of potential primary paths. For each potential primary path, the remainder of the set comprises its associated set of potential backup paths from which a designated backup path for that potential primary path is chosen. The designated backup path is selected based on channel sharing opportunity as estimated using our inventive method. For each potential primary path, our method uses the limited link information found locally at each network element to estimate for each link in each potential backup path the likelihood that the primary path under consideration is failure risk-diverse to the set of primary paths already being protected by that link, and thus the probability that a channel in this link can be shared. The estimate of the ability of the backup links to be shared is used to weight the cost associated with each link in the backup paths, which is further used to select the lowest cost backup path for each potential primary path. The resulting set of potential primary paths and their associated backup paths are then ordered from lowest to highest estimated cost. Signaling between the ingress network element and each network element in the proposed paths is then used to provision the lowest estimated cost path pair. If the lowest estimated cost pair can not be established, it is removed from the pool of candidates and the next lowest estimated cost pair is tried. This process continues until either a successful attempt is made or the pool of candidates is exhausted. If the candidate pool becomes exhausted, an error signal is returned to the network controller and the attempts cease.

[021] Viewed from another aspect, our invention is directed to communications networks employing a stochastic method of path computation in order to reduce the amount of resources required to provide desired services. The provisioning of network paths according our inventive method can occur at any time, such as predetermined intervals, or at every new path provisioning event in order to attain and maintain efficient use of network resources.

[022] **Brief Description of the Drawings**

[023] Figure 1a shows a shared mesh-restoration network with two established primary and backup path pairs;

[024] Figure 1b shows the network of Figure 1a after the provisioning of a new primary and backup path pair;

[025] Figure 2 shows the steps involved in determining the probability that a channel in a backup path link can be shared with a proposed backup path;

[026] Figure 3 shows the steps involved in provisioning a path pair in connection with the probability method described in Figure 2 according to the present invention;

[027] Figure 4 is a plot of the error distribution of the estimate probabilities minus experimental probabilities according to our inventive method for a large number of random topographical arrangements;

[028] Figure 5 is a plot of the error distribution of the exact probabilities obtained by computation according to our inventive method for a large number of random topographical arrangements;

[029] Figure 6 shows a summary of the results comparing the stochastic method according to our inventive method with a deterministic method as implemented in a centralized control topology;

[030] Figure 7 shows the distribution of sharing probabilities in a 100-node, 137-link network; and

[031] Figure 8 shows the distribution of sharing probabilities in a 220-node, 300-link network.

[032] **Detailed Description of the Invention**

[033] Our inventive method provides a cost-efficient means of making a connection between communications equipment connected to an ingress network element and communications

equipment connected to an egress network element through a shared-mesh protection network utilizing distributed network control.

[034] In order to better illustrate our invention, Figure 1a depicts a mesh-connected network prior to provisioning new service and Figure 1b depicts the same network after provisioning of new service according to our inventive method. In Figure 1a, a primary path, 160-1, carries a bi-directional signal between a client equipment 120-1 connected to network element 110-1 and client equipment 120-2 connected to network element 110-2. Primary path 160-1 includes network element 110-1, link 130-1, and network element 110-2, and traverses shared-risk group (SRG) 150-1. Primary path 160-1 has an associated backup path 170-1, having network elements 110-1, 110-3, 110-4, 110-2, and links 130-2, 130-4, and 130-3, traversing SRGs 150-5, 150-2, 150-3, 150-4, and 150-6.

[035] A second primary path, 160-2, carries a signal between client equipment 120-4 connected to network element 110-5 and client equipment 120-3 connected to network element 110-4. Primary path 160-2 is comprised of network element 110-5, link 130-7, and network element 110-4, and its link traverses SRGs 150-10 and 150-11. Separate SRGs 150-10 and 150-11 in link 130-7 are shown to demonstrate a commonly occurring situation wherein multiple SRGs relate to a single link, for example a transmission line travelling through multiple conduits. Primary path 160-2 has an associated backup path 170-2, which includes network elements 110-5, 110-3 110-4, and links 130-6 and 130-4, traversing SRGs 150-7, 150-2, 150-3, and 150-4.

[036] Primary paths 160-1 and 160-2 are the shortest paths that achieve desired connectivity, requiring the least number of network elements and links and thus incurring the lowest cost.

[037] Since the primary paths shown in Figure 1a have no common SRGs, their respective backup paths are allowed to share equipment if desirable. In the case of backup paths 170-1 and 170-2, the shortest potential backup paths for each would include link 130-4, so they are shown sharing the resources of network elements 110-3, 110-4, and link 130-4.

[038] Figure 1b shows the provisioning of new service between client equipment 120-4 and 120-5 according to our preferred embodiment. A request for new service is generated by client equipment 120-4 and received by 110-5, the ingress network element for this service. The ingress node accesses a locally available database of state variables to develop a set of K shortest paths through the network interconnecting network elements 110-4 and 110-5. In this case, the maximum available value of K is equal to 4 and a complete set of potential primary paths can be denoted by the following list of paths w_i: where w_1={130-8}, w_2={130-7,130-9}, w_3={130-6,130-4, 130-9), and w_4={130-6,130-2,130-1,130-3,130-9}.

[039] For each potential primary path in the set, proceeding from shortest to longest, a potential backup path is chosen from the remainder of the set according to our inventive method which employs a weighting scheme with the purpose of selecting the lowest weight backup path for each primary path. The weighting scheme assigns weighted costs to links in potential backup paths in a manner that encourages sharing of channels among multiple primary paths. Links contained in a prospective primary path or that are not SRG-disjoint with the prospective primary path are assigned an infinite weight. Links that do not have an available shared protection channel are assigned a weight corresponding to the real cost of adding a shared protection channel to that link. Links that have an available shared protection channel are assigned a weighted cost corresponding to the actual cost of that channel scaled by the probability that the channel can not be shared. Thus, the cost of a link that has a high sharing probability would be multiplied by a very low weight, giving it very low weighted cost. A link that has high probability of SRG contention with a primary path already being protected by that link would be assigned a weighted cost very close to its actual cost. The total cost of the paths are then estimated by adding the weighted costs of all the links contained therein.

[040] In the example shown here, link 130-7 is assigned an infinite weighted cost due to a shared SRG 150-10 with the shortest primary path 160-3, and link 130-4 is assigned an infinite weighted cost due to its previously provisioned use as a backup path for primary path 160-2 that shares SRG 150-10 with potential primary path 150-3. As a result, paths w_2 and

w_3 are assigned an infinite cost, removing them from consideration and leaving only w_1 and w_4 as viable paths for this connection.

[041] Beginning with the path pair containing the shorter primary path $\{w_1,w_4\}$, the probability of allowed sharing of links in the backup path is computed using their SRG information known by network element 110-5 and a weighted cost of the path pair is determined. The process is repeated for all path pairs resulting in weighted cost estimates for all potential path pairs.

[042] The weighted cost estimates are used to define a set of path pairs comprised of each potential primary path and its corresponding shortest, i.e. lowest cost, backup path. After ordering the path pairs from lowest cost, $\{w_1,w_4\}$, to highest cost, $\{ w_4,w_1 \}$, signaling between the ingress network element, 110-5, and all other network elements in the paths is used to attempt to provision the lowest cost path pair. If any link in either w_1 or w_4 can not be established, $\{w_1,w_4\}$ is removed from consideration and the ingress network element attempts to provision the new lowest cost path pair, $\{w_4,w_1\}$. If the ingress element can not establish any of the path pairs in the set of candidates, an error message is transmitted to the client equipment that requested the service.

[043] In one embodiment, our inventive method is used to compute a probability that a shared protection channel in a backup link can be shared as described in Figure 2. The probability is computed by determining the number of available shared protection channels existing in the link and assigning this value to M, as in step 210, creating a set of N SRGs traversed by the primary to be protected and labeling them as SRG_j where $1 \leq j \leq N$, as in step 220, assigning the value of the number of times SRG_j is protected by the shared protection channel set in the link to n_j, as in step 230, computing the probability, p, that any one shared protection channel is shareable as $p=\prod_j(1-n_j/M)$, as in step 240, and computing the probability, P, that at least one shared protection channel is shareable as $P=1-(1-\prod_j(1-n_j/M))^M$, as in step 250.

[044] In another embodiment of our invention, a graphical method of provisioning primary and backup paths is used as described in Figure 3. The steps of our inventive method include

creating a graph G(V,E) wherein each network element in the network is represented by a vertex, V, each bi-directional link between adjacent network elements is represented by an edge, E, and a source vertex corresponding to the ingress network element and a destination vertex corresponding to the egress network element are defined as in step 310. Further steps include calculating a set of K shortest paths from ingress network element to egress network element through the network, ordered by length, and designated as potential primary paths w_i, where i ranges from 1 to K as in step 320, initializing the set of candidate path pairs, S, equal to the null set, $\varnothing$, as in step 330, determining a second shortest path corresponding to each first shortest path, w_i, between ingress network element and egress network element to be designated as the protection path, s_i, and returning a candidate pair of paths, $\{w_i,s_i\}$, as in step 340.

[045] The candidate pair of paths is determined by setting the weight for each edge that shares an SRG with w_i, ∞, as in step 341, setting the weight for each edge that has neither shared protection channel nor unassigned primary channel to ∞, as in step 342, setting the weight of each edge without a shared protection channel to the actual cost of the edge, as in step 343, using the probability method shown in Figure 2, as in step 344, setting the weight of each edge with at least one shared protection channel to the real cost of the edge scaled by the probability that no shared protection channel is shareable as determined in as in steps 210 to 250 of Figure 2, as in step 345, calculating the shortest potential backup path associated with w_i and assigning it to path s_i, as in step 346, adding the minimum weight path pair {wi,si} to the set of potential path pair candidates, S, as in step 350, selecting the minimum weight pair $\{w_m,s_m\}$ from the set S, as in step 360, using signaling between the ingress network element and all other network elements in w_m and s_m to establish the links comprising the paths, eliminating this pair from S and returning to Step 360 to select a new minimum weight pair $\{w_m,s_m\}$ from set S if any link in the paths can not be provisioned, continuing until primary and backup paths are successfully provisioned, or S is exhausted, and returning an error signal to the client equipment requesting service if S becomes exhausted, as in step 370.

[046] Experimental verification of our inventive method through simulation and experimentation on an optical network is shown in Figure 4, Figure 5, Figure 6, Figure 7, and Figure 8. Qualitatively, the complexity of our stochastic method is the product of the number of SRGs traversed by the primary path, N, and the probability of any single path being sharable, p, raised to the power of the number of shared protection channels in a link, M. Note that the complexity is independent of the number of paths in the entire network, and the computation can be realized in O(N+logM) time.

[047] The results obtained with our stochastic method and distributed control are in sharp contrast to those obtained using a deterministic method and centralized control in that the latter method has both time and space complexity dependent upon the number of paths in the network and therefore has difficulty scaling as the network grows.

[048] The quality of the estimated probability that a link contains a shareable shared protection channel given the information on the number of times each SRG traversed by the primary path is restored in that link can be experimentally verified. The experiment consists of simulating several millions of random arrangements, and compute the ratio of combinations with available shared protection channels to the number of combinations without available shared protection channels. The difference between each experimental probability and the corresponding exact and approximate probabilities obtained by computation is then compared. Figure 4 shows the error distribution of the estimate probabilities minus experimental probabilities obtained over the range of problem instances. Figure 5 shows the error distribution of the exact probabilities obtained through direct computation. We observe that the estimate probability has a tendency to underestimate the experimental probability, but it is accurate within 0.05 for 85% of the time, which quite remarkable given the simplicity of its computation. In comparison, the simulation exhibits an accuracy within 0.01 of the exact probability, and a closer look indicates that 70% of the time the difference is within $5x10^{-4}$.

[049] As can be readily appreciated by those skilled in the art, our invention advantageously produces highly efficient network usage such as that computed using a centralized control

scheme, but with much faster computation times associated with distributed control architectures. As evidence of the computation time advantage, we consider two scenarios inspired from real life networks. NetA is a 100- node, 137-edge network, with one unit of demand between every pair of node (4950 demands). NetB is 220-node, 300-edge network, also with one unit of demand between every pair of node (24090 demands.) For the sake of simplicity we assume that every edge costs one unit of currency and corresponds to one SRG (i.e. one SRG per edge and one edge per SRG). We then route the demands on each network using the deterministic and the stochastic methods. We are interested here in the processing time to complete each method, and the quality of the solutions expressed in total number of channels required (used for primaries and reserved for backups.) The results are summarized in Figure 6. For NetA (resp. NetB) we observe that the stochastic approach is 6.78 time faster (resp. 19.7 time faster) than the deterministic approach while the penalty is only 2% (resp. 3%) more capacity. Also important is the amount of information the route computation module (RCM) needs to compute the routes. The stochastic based RCM only require one array per edge, where each entry indicates the number of times the SRG is protected in the edge by any reserved channel. In the NetB problem they are thus 300 such arrays (one per edge) of 300 entries each (one per SRG). For comparison, the deterministic approach needs an array for each reserved channel, where each entry corresponds to an SRG and indicates whether the SRG is protected or not by the reserved channel. In NetB 213052 of the channels are reserved for protection, thus 213052 arrays of 300 entries each are required.

[050] Figures 7 and 8 show the distribution of sharing probabilities as computed by the stochastic method during the provisioning of the demand in NetA and NetB. The distribution in Figure 7 shows that the probability of having a sharable channel was essentially zero (<0.05) 48% of the time and essentially 1.0 for over 20% of the time. Therefore, it was possible to determine with near certainty the availability of a sharable reserved channel in NetA approximately 70% of the time. Figure 8 shows somewhat improved results with near certainty approximately 77% of the time.

[051] It will be understood that the embodiments of the present invention specifically shown and described are merely exemplary and that a person skilled in the art can make alternate embodiments using different configurations and functionally equivalent components. All such alternate embodiments are intended to be included in the scope of this invention as set forth in the following claims.